AWARD NUMBER: W81XWH-13-1-0088

TITLE: Exploring the Presence of microDNAs in Prostate Cancer Cell Lines, Tissue, and Sera of Prostate Cancer Patients and its Possible Application as Biomarker

PRINCIPAL INVESTIGATOR: Dr. PANKAJ KUMAR

CONTRACTING ORGANIZATION: RECTOR & VISITORS OF THE UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VA 22903

REPORT DATE: August 2014

TYPE OF REPORT: Annual report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
August 2014	Annual	1 Aug 2013 – 31 Jul 2014

4. TITLE AND SUBTITLE

Exploring the Presence of microDNAs in Prostate Cancer Cell Lines, Tissue, and Sera of Prostate Cancer Patients and its Possible Application as Biomarker

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-13-1-0088

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Pankaj Kumar

E-Mail: pk7z@virginia.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
RECTOR & VISITORS OF THE UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VA 22903

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

MicroDNAs are extra chromosomal circular DNA present in normal mammalian somatic cells. To find the prostate tissue-specific microDNA a panel of human prostrate (LnCap (PSA, hK2 and AR positive), C4-2, and PC-3 (non-tranformed prostate epithelium)) and ovarian (ES2 and OVCAR-8) cancer cell lines were examined for microDNA. The identified microDNAs in all the cell lines were mostly ~200bp or ~400bp in size, arising from the GC rich regions in the genome, and mostly mapped to genic regions and are not associated with repetitive DNA sequences. MicroDNA are enriched in area of genome that had high exon density suggesting a role of splicing in microDNA generation. Comparison of microDNA loci across the cell lines identified hot spots of microDNA generation that are present on every chromosome, and correspond to areas of high gene density and high GC content. However, hierarchical clustering on the basis of microDNA co-ordinates classified the prostate and ovarian cancer cell lines into two separate groups suggesting that at least some microDNAs are tissue-specific and so their sites of origin are affected by tissue-specific gene expression patterns or epigenetic marks.

15. SUBJECT TERMS
microDNA; eccDNA; Prostate Cancer; Biomarker

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	9	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

INTRODUCTION:

Along with colleagues, I have recently discovered a new type of extra-chromosomal circular DNA (eccDNAs, also called microDNA) in mouse tissue as well as in mouse and human cell lines (*1*). These eccDNAs arise from tens of thousands of unique genomic loci and could serve as disease biomarkers. Discovering a new biomarker for prostate cancer is significant because early detection and accurate prognosis is very important to cure the disease without over treating the many patients who do not have life-threatening disease. In this current project I am exploring the potential of microDNA as prostate cancer biomarker. The circular DNAs will be stable due to their circular nature (resistant to exo-nuclease) and could be amplified by PCR based method. Finally I will look for the presence of prostate cancer specific circular DNA in sera of prostate cancer patients.

KEYWORDS: eccDNA; microDNA; high-through put sequencing; prostate cancer; sera; biomarker

OVERALL PROJECT SUMMARY: High-throughput sequencing of extra chromosomal circular DNA (eccDNA): Major Task 1 (1-8 months):

Isolation and high-throughput sequencing of eccDNA from prostate and non-prostate derived cell lines

Extraction of circular DNA: The steps involved in the isolation of circular DNA are shown in **Fig. 1**. In brief, the nuclei from the cells were extracted as described (Shibata et al. 2012). To avoid contamination by mitochondrial DNA only the nuclei of the cell lines were used for the extraction of eccDNA (Jiang et al. 2008; Shibata et al. 2012). Contaminating linear DNA was removed by an ATP-dependent exonuclease (*1, 2*). Purified extra-chromosomal fraction was treated sequentially with proteinase K and RNase, with phenol-chloroform extraction and ethanol precipitation. Multiple displacement amplification (MDA) with random hexamers (*1, 3, 4*) was used to enrich circular DNA by rolling circle amplification. This procedure was applied to isolate eccDNA from three prostate cell lines: LNCaP (PSA, hK2 and AR positive), PC-3 & C4-2, (non-tranformed prostate epithelium) and two ovarian cancer cell lines (ES2 and OVCAR-8). The summary of isolation of microDNA and its yield is shown in Table 1.

Table 1: Summary of microDNA isolation in various cancer cell lines.

Cell Line	**ES2**	**OVCAR8**	**LnCap**	**PC-3**	**C4-2**
Cell Count	$1.8X10^8$	$1x10^8$	$1.1X10^8$	$1.24X10^8$	$1.1X10^8$
Episomal DNA (ug)	21.3	23.7	26	15.6	20.4
Starting DNA (ug)	21.3	23.7	26	10	20
ExoVII (ng)	5600	9632	6074	2640	9352
ATP-dependent DNase (ng)	350	530	680	466	326
Rolling Circle Amplification (RCA) Starting DNA (ng)	88	133	120	116.5	81.5
RCA Ending DNA (ug)	9.2	4	8.368	6.8	8.515
DNA Shearing (400-600bp) (ng)	1472	1312	936	1116	1090
MicroDNA Library (ng)	402	423	738	528	1224
MicroDNA Library conc (ng/uL)	13.4	14.1	24.8	17.6	40.8

MicroDNA library preparation and sequencing: Enriched eccDNA was fragmented, selected and sequenced (Sanger sequencing) to verify the presence of circular DNA. Cloning and sequencing of 500 bps long MDA product confirmed circular nature of DNA (**Fig. 2**). Once circular nature of DNA was confirmed

then paired-end (PE) library was prepared as described (*1*). The 500bp size selection was done on nebulized DNA. The ends of the library fragments were modified as per Illumina paired-end protocol and paired-end high-throughput sequencing (64 bases long reads) was performed according to the manufacturer's protocol (Illumina). Summary of microDNA sequencing and mapping in prostate and ovarian cancer cell lines is shown in Table 2.

Table 2: Summary of PE sequencing and mapping to human genome.

Sample	Paired End Reads	Pairs Aligned	Read Sequences	Aligned Sequences	Unique Alignment	Unique microDNA (Complexity)
ES2	61.9	26.8	123.9	96.4	86.7	**114,752**
OVCAR8	50.2	28.8	100.4	84.5	75.8	**57,327**
C4-2	41.1	21.4	82.2	69.3	63.2	**41,410**
LnCap	56.1	24.8	112.1	89.1	82.3	**84,841**
PC3	43.5	10.7	87.0	41.6	38.8	**14,705**

*all values in millions except microDNA

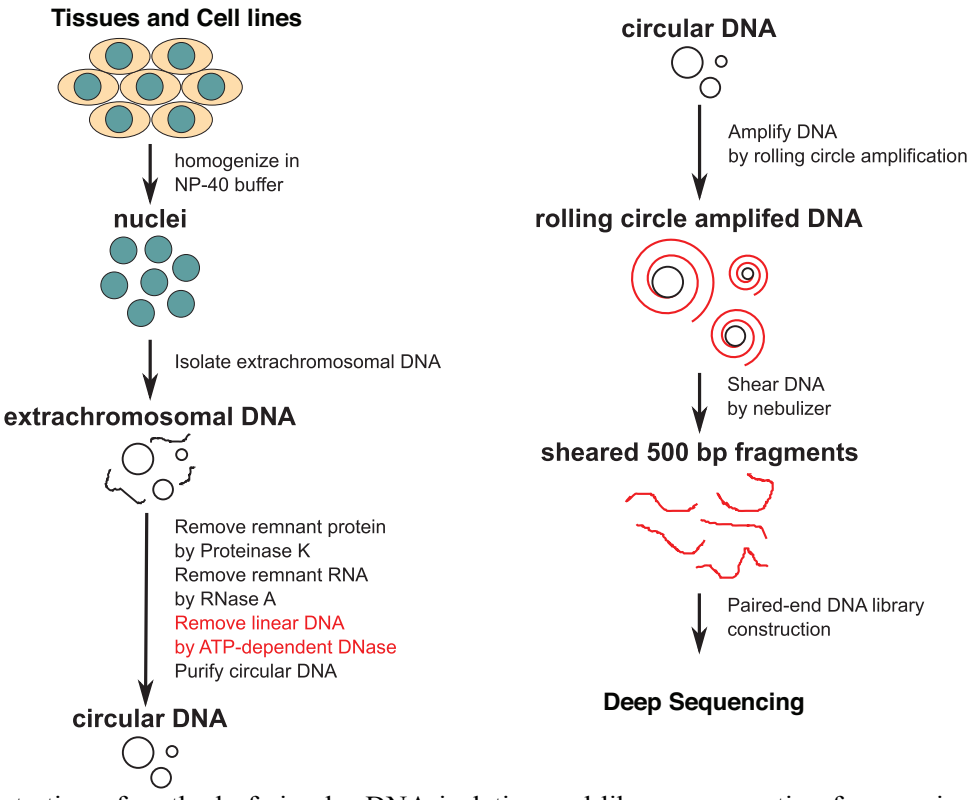

Figure 1: Illustration of method of circular DNA isolation and library preparation from various cell lines of prostate and ovarian tissue. ATP-dependent DNase-resistant DNA from nuclei (eccDNA) was amplified by multiple-displacement amplification (MDA). The amplified DNA was sheared to obtain 500 bp fragments and sequenced by the Illumina sequencing.

GCAGCACCATTTACAATGATGCTGCACATTAAATTCAACAGGGAGAAATCCTCTCTGCCCCTCAG
ACTGCCCATCAGGCTTGGGAGGTGTCGGGAGACAGGCGTTCATCCTGGTCGCTGCTTTGGGTAG
CAGCTTGCAGTGCTGAAACAGTCAAAGATGGCTGTCCCTCAGCCCTGCCACCTCCCATTCAAGCG
CCTGCTCTGAAAGCTCCTGAGCAGATGGGCCTGAGATGCAGACAGGGGTGCTCGTGGCAGCACC
ATTTACAATGATGCTGCACATTAAATTCAACAGGGAGAAATCCTCTCTGCCCCTCAGACTGCCCA
TCAGGCTTGGGAGGTGTCGGGAGACAGGCGTTCATCCTGGTCGCTGCTTTGGGTAGCAGCTTGC
AGTGCTGAAACAGTCAAAGATGGCTGTCCCTC

Figure 2: Presence of repeat sequence in the 500 bps long cloned and Sanger sequenced fragment. Circular DNA sequence is in purple.

Major Task 2: Identification of circular DNA from various samples (9-15 months)

MicroDNA identification from the paired-end sequencing: The details of the different steps of identification of microDNA are shown in **Fig. 3**. Sequence tags were mapped on the human reference genome using the Novoalign software. Only those tags that were mapped uniquely were considered for the identification of circular DNA. The sequence coverage of each base pair was profiled for each chromosome. An island of interest (potential circle) was delineated on the basis of two consecutive sequenced bases. In other words any stretch of continuously sequenced bases was considered as a part of an island and the start and end of the stretch was considered as start and end of the island respectively. The islands were considered further for the identification of circles. The creation of circular microDNAs would bring together the ends of the linear islands to create a novel junctional sequence that does not exist in the genome. Thus the PE-sequence of a fragment that breaks at or very close to a junction will have one end that maps to the island and another end that maps to the junction and will not map to the reference genome (**Fig. 3b**). Those PE-tags where one tag maps uniquely to an island and the other remains unmapped, but passes the sequence quality filter, was considered for the validation of circular nature of the identified islands. As mentioned earlier, the creation of circular DNAs bring together the two ends of the linear DNA, and thus generated hypothetical junctional tags was created by ligation of the two ends of each island. If the mapped tag of a PE read falls in an island and the un-mapped tag matches a hypothetical junctional tag of the same island, then the island was annotated as a circle. Summary of microDNA sequencing and mapping and number of microDNA identified in prostate and ovarian cancer cell lines is given in **Table 2**.

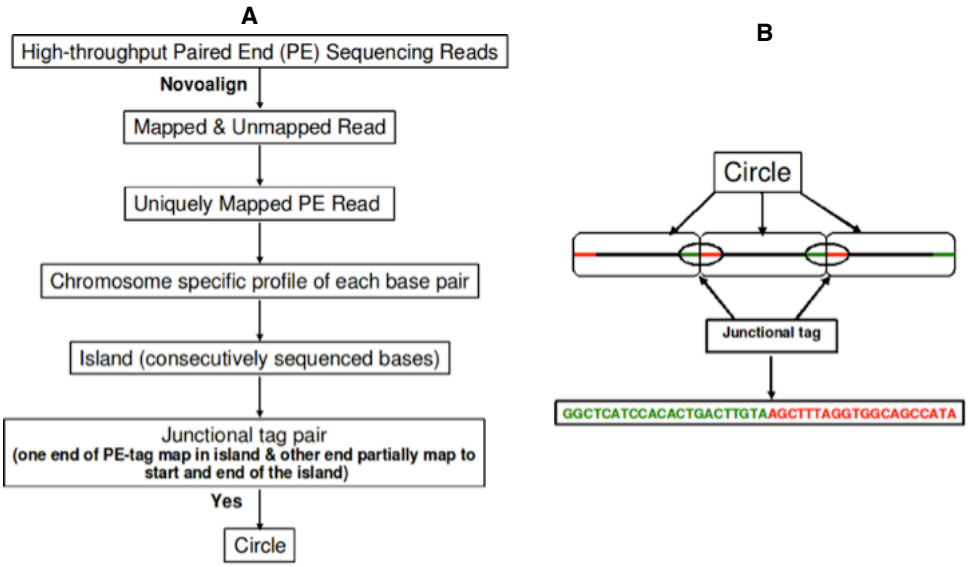

Figure 3: Illustration of different steps in the identification of microDNA by Island method (a) and schematic representation of junctional tag (b).

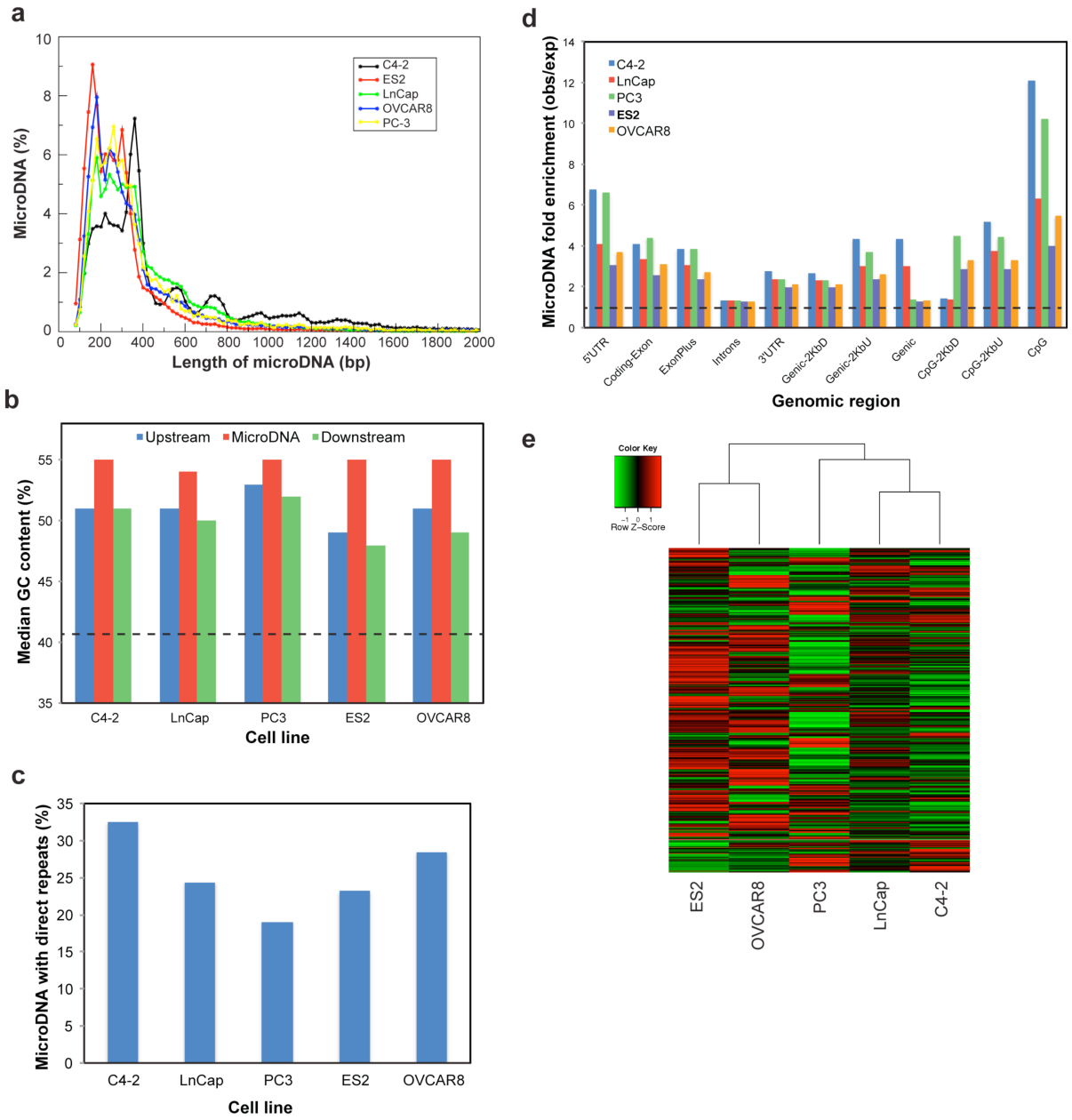

Figure 4: Properties of microDNA identified in human prostate and ovarian cancer cell lines. (a) Length distribution of microDNAs identified in cancer cell lines. (b) Median percent GC content of microDNAs and the genomic sequences up- or downstream of the source loci are enriched relative to the average GC content of the human genome (dashed line). (c) Direct repeats near the start and end of microDNA sequences (2- to 15-bp) are enriched in all cell lines compared to a random model (RM). (d) Enrichment of microDNAs in the indicated genomic region relative to the expected percentage based on random distribution. (e) MicroDNA loci were grouped into 5-Mb bins stepwise across the human genome and the percentage of all microDNA located within each bin was calculated for each cancer cell line and compared using hierarchical clustering.

Identification and analysis of prostate-tissue-specific and prostate-cancer-specific microDNAs:

First, all microDNAs were studied for general properties: size distribution (**Fig. 4a**), GC content (**Fig. 4b**), the presence of 2-10 base direct repeats at the ends of microDNA (**Fig. 4c**), and locations relative to genomic features (exons, introns, UTRs, CpG islands) (**Fig. 4d**). It could be seen that most of the microDNAs are of 200-400 bps long, have high GC content compared to the genomic average and frequently have direct repeat at the ends. These features are similar to the features that have all been observed in the microDNAs identified in normal mouse tissue, and mouse NIH3T3 and human HeLa cells (*1*). This confirms that the microDNAs obtained from normal tissue and human cell lines of different tissue origin conforms to the general properties of microDNAs.

The genomic origins of the circles and the abundance of circles from each locus were compared by hierarchical clustering. For this whole human genome was divided in 5-mega base windows and in each bin the fraction of microDNA in each bin was calculated and compared using hierarchical clustering. It is interesting to note that prostate cancer cell lines are clustering together (**Fig. 4e**) and distinct from the ovarian cell line cluster indicating that some of the genomic loci are differentially producing microDNA between prostate and ovarian cell line. The common and abundant circles identified across all the prostate cancer cell lines have the potential to qualify as a marker for prostate cancer however this tissue specificity of microDNA need to be further checked in patient sera.

In the next level of my study I propose to look for the presence of prostate cancer specific circular DNA in sera of prostate cancer patients. Even the identification of microDNAs in serum will be a novel discovery.

KEY RESEARCH ACCOMPLISHMENTS:

- MicroDNAs are present in cancer cell lines
- MicroDNAs identified in cancer cell lines have similar features (length distribution, GC content, genomic enrichment etc.) that have been observed in the microDNAs identified in normal mouse tissue, and mouse NIH3T3 and human HeLa cells
- Hierarchical clustering of prostate and ovarian cancer cell lines based on the microDNA loci in the genome indicates some of the microDNAs are tissue specific
- Tissue specific microDNA could be disease biomarker

CONCLUSION:

To find the tissue specific microDNA we examined a panel of human prostrate (C4-2, LnCap and PC-3) and ovarian (ES2 and OVCAR-8) cancer cell lines. Hierarchical clustering on the basis of microDNA co-ordinates classified the prostate and ovarian cancer cell lines into two separate groups suggesting that microDNA are tissue specific. The tissue specificity of these microDNA could be further explored to find prostate tumor specific microDNAs that could serve as biomarkers for cancer detection and its prognosis. DNA, especially circular DNA, is extremely stable and is also expected to survive in the blood once it is released from cancer cells.

In the last level of my study I will look for the presence of prostate cancer specific circular DNA in sera of prostate cancer patients. Even the identification of microDNAs in serum will be a novel discovery. We will also test whether the sequences, abundance, and nature of the microDNAs are different in the normal sera compared to sera from the four prostate cancer patients. The preliminary data from this part of the project will be essential before we can propose a more definitive project on these lines.

PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS:

a. List all manuscripts submitted for publication during the period covered by this report resulting from this project.

1. Lay Press: "Nothing to report."
2. Peer-Reviewed Scientific Journals:

Laura W. Dillon,[1]* **Pankaj Kumar**,[1]* Yoshiyuki Shibata,[1]* Teressa M. Paulsen,[1] Smaranda Willcox,[2] Jack D. Griffith,[2] Yves Pommier,[3] Shunichi Takeda,[4] Anindya Dutta[1] **Extrachromosomal microDNAs in vertebrate tissues arise from areas with active chromatin marks and high exon density by mismatch repair and other repair pathways** (manuscript under preparation). * Joint first author

[1]Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA, USA
[2]Lineberger Cancer Center, University of North Carolina, Chapel Hill, NC, USA
[3]Laboratory of Molecular Pharmacology and Developmental Therapeutics Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892-4255, USA
[4]CREST Research Project, Japan Science and Technology Corporation, Radiation Genetics, Faculty of Medicine, Kyoto University, Konoe Yoshida, Sakyo-ku, Kyoto 606-8501, Japan

3. Invited Articles: "Nothing to report."
4. **Abstracts:**

P Kumar, Laura W. Dillon, Y Shibata, and A Dutta., **MicroDNA (Extra Chromosomal Circular DNA) in Normal Mammalian Tissues and Cancer Cell Lines**; (Accepted for poster presentation at the 64th Annual Meeting of The American Society of Human Genetics, October 20, 2014 in San Diego, CA.

b. List presentations made during the last year (international, national, local societies, military meetings, etc.). "Nothing to report."

INVENTIONS, PATENTS AND LICENSES: "Nothing to report."

REPORTABLE OUTCOMES: "Nothing to report."

OTHER ACHIEVEMENTS: "Nothing to report."

REFERENCES:

1. **Y. Shibata *et al.*, Extrachromosomal MicroDNAs and Chromosomal Microdeletions in Normal Tissues. *Science* 336, 82 (Apr 6, 2012).**
2. **H. Yamagishi *et al.*, Purification of Small Polydisperse Circular DNA of Eukaryotic Cells by Use of Atp-Dependent Deoxyribonuclease. *Gene* 26, 317 (1983).**
3. **F. B. Dean *et al.*, Comprehensive human genome amplification using multiple displacement amplification. *P Natl Acad Sci USA* 99, 5261 (Apr 16, 2002).**
4. **L. Lovmar, A. C. Syvanen, Multiple displacement amplification to create a long-lasting source of DNA for genetic studies. *Hum Mutat* 27, 603 (Jul, 2006).**

APPENDICES: none